

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKewed/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-331847

(43) 公開日 平成11年(1999)11月30日

(51) Int.Cl.⁶

識別記号

F I

H 0 4 N 7/30

H 0 4 N 7/133

Z

1/41

1/41

B

審査請求 未請求 請求項の数39 O L (全 18 頁)

(21) 出願番号 特願平11-89985

(22) 出願日 平成11年(1999) 3 月30日

(31) 優先権主張番号 0 9 / 0 5 2 , 0 3 9

(32) 優先日 1998年 3 月30日

(33) 優先権主張国 米国 (U S)

(71) 出願人 000002369

セイコーエプソン株式会社
東京都新宿区西新宿 2 丁目 4 番 1 号(72) 発明者 ヴィレッシュ ラトナカー
アメリカ合衆国カリフォルニア州94086
サニーベイル コルトマデラアベニュー
970 アパートメント402(72) 発明者 ビクター アイバシン
アメリカ合衆国カリフォルニア州94506
ダンビル サットンコート19

(74) 代理人 弁理士 鈴木 喜三郎 (外 2 名)

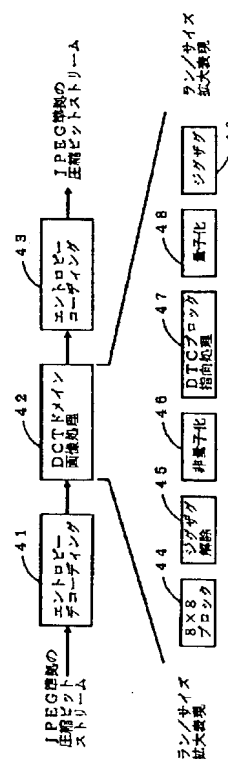
最終頁に続く

(54) 【発明の名称】 画像変換方法、ディジタルカメラおよびコンピュータシステム

(57) 【要約】

【課題】 完全な伸張および圧縮の工程を経由すること無しに、画像の圧縮ドメイン表現を直接操作して、選択された空間ドメイン操作を行う。

【解決手段】 完全な伸張および圧縮工程を行う必要なしに、所望の空間ドメイン処理を達成するための、画像の圧縮ドメイン表現の操作を含む画像処理法。この手法には、エントロピーデコーディング41の後に、DCTドメイン画像処理42を行う。このDCTドメイン画像処理42には、J P E G 画像上のD₄ (正方形の対称の二面グループ) における8個のオペレーションを行うためのアプローチの処理が含まれる。画像の90°回転(D₄でのオペレーション)のようなタスクについては、DCTドメインに基づく方法により、空間ドメインに基づく手法よりも速度がほぼ5倍増となる。これらの簡易な圧縮ドメインに基づく処理法は、J P E G に基づくスチルカメラシステムに必要な画像処理タスクに十分適合する。



【特許請求の範囲】

【請求項1】 デジタル画像の一次変換ドメイン表現を操作することにより、デジタル画像の空間ドメイン表現上で二面对称オペレーションを行う画像変換方法であって、

前記デジタル画像の一次変換ドメイン表現を規定する、複数の一次変換に基づくデータブロックを抽出するステップ、

前記複数の一次変換に基づくデータブロックを再配列するステップ、

前記複数の一次変換に基づくデータブロックのうちの少なくとも1つに一次変換ドメイン・オペレーションを適用するステップ、および前記複数の一次変換に基づくデータブロックを再アッセンブル化するステップ、を含んでなり、前記デジタル画像の空間ドメイン表現が二面对称オペレーションに付される、ことを特徴とする画像変換方法。

【請求項2】 圧縮画像データを解析して、前記複数の一次変換に基づく各データブロックについてオフセット値およびDC値を抽出するステップ、

対応するオフセット値を用いることにより前記複数の一次変換に基づく各データブロックをアクセスするステップ、および前記複数の一次変換に基づく各データブロックの要素を抽出するステップ、をさらに含む請求項1記載の画像変換方法。

【請求項3】 前記二面对称オペレーションが主対角線上でデジタル画像をフリップすることを含む請求項1または2に記載の画像変換方法。

【請求項4】 前記二面对称オペレーションが中央垂直軸上でデジタル画像をフリップすることを含む請求項1または2に記載の画像変換方法。

【請求項5】 前記二面对称オペレーションが交差方向対角線上でデジタル画像をフリップすることを含む請求項1または2に記載の画像変換方法。

【請求項6】 前記二面对称オペレーションが中央水平軸線上でデジタル画像をフリップすることを含む請求項1または2に記載の画像変換方法。

【請求項7】 前記二面对称オペレーションがデジタル画像を時計回り方向に90°回転させることを含む請求項1または2に記載の画像変換方法。

【請求項8】 前記二面对称オペレーションがデジタル画像を180°回転させることを含む請求項1または2に記載の画像変換方法。

【請求項9】 前記二面对称オペレーションがデジタル画像を反時計回り方向に90°回転させることを含む請求項1または2に記載の画像変換方法。

【請求項10】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内で要素を転置することを含む請求項3に記載の画像変換方法。

【請求項11】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内で奇数列要素の符号を反転させることを含む請求項4に記載の画像変換方法。

【請求項12】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内で要素を転置し、次いで各ブロック内のその他あらゆる要素の符号を反転させることを含む請求項5に記載の画像変換方法。

【請求項13】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内で奇数行要素の符号を反転させることを含む請求項6に記載の画像変換方法。

【請求項14】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内で要素を転置し、次いで各ブロック内の奇数列要素の符号を反転させることを含む請求項7に記載の画像変換方法。

【請求項15】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内のその他あらゆる要素の符号を反転させることを含む請求項8に記載の画像変換方法。

【請求項16】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内で要素を転置し、次いで各ブロック内の奇数行要素の符号を反転させることを含む請求項9に記載の画像変換方法。

【請求項17】 光を捕捉し、光をアナログ画像信号に変換するためのセンサーと、

空間ドメイン内でアナログ画像信号をデジタル画像に変換するためのデジタル・アナログ変換器と、

デジタル画像の空間ドメイン表現から、複数の一次変換に基づくデータブロックで規定されるデジタル画像の一次変換ドメイン表現に、デジタル画像を圧縮するためのデジタル画像プロセッサと、を含んでなり、前記デジタル画像プロセッサは、

前記複数の一次変換に基づくデータブロックを抽出し、当該複数の一次変換に基づくデータブロックを再配列し、

前記複数の一次変換に基づくデータブロックのうち少なくとも1つに一次変換ドメイン・オペレーションを適用し、前記複数の一次変換に基づくデータブロックを再アッセンブル化してデジタル画像の空間ドメイン表現上で二面对称オペレーションを実行する、ことを特徴とするデジタルカメラ。

【請求項18】 前記二面对称オペレーションがデジタル画像を時計回りに90°回転させることを含む請求項17に記載のデジタルカメラ。

【請求項19】 前記二面对称オペレーションがデジタル画像を180°回転させることを含む請求項17に記載のデジタルカメラ。

【請求項20】 前記二面对称オペレーションがデジタル画像を反時計回りに90°回転させることを含む請求項17に記載のデジタルカメラ。

【請求項21】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内でエレメントを転置し、次いで各ブロック内の奇数列エレメントの符号を反転させる請求項18に記載のデジタルカメラ。

【請求項22】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内のその他あらゆるエレメントの符号を反転させる請求項19に記載のデジタルカメラ。

【請求項23】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内でエレメントを転置し、次いで各ブロック内の奇数行エレメントの符号を反転させる請求項20に記載のデジタルカメラ。

【請求項24】 コンピュータシステムに空間ドメイン内でデジタル画像の幾何学的変換を生じさせるために組み込まれたコンピュータ可読プログラムコードを有するコンピュータシステムであって、前記デジタル画像を処理して、複数の一次変換に基づくデータブロックで規定されるデジタル画像の一次変換ドメイン表現を得るステップ、前記複数の一次変換に基づくデータブロックを抽出するステップ、前記複数の一次変換に基づくデータブロックを再配列するステップ、前記複数の一次変換に基づくデータブロックのうちの少なくとも1つに一次変換ドメイン・オペレーションを適用するステップ、および、前記複数の一次変換に基づくデータブロックを再アセンブル化するステップ、を含んでなり、デジタル画像の空間ドメイン表現が二面对称オペレーションに付されることを特徴とするコンピュータシステム。

【請求項25】 圧縮画像データを解析して、前記複数の一次変換に基づく各データブロックについてオフセット値およびDC値を抽出するステップ、対応するオフセット値を用いることにより、前記複数の一次変換に基づく各データブロックをアクセスするステップ、および、前記複数の一次変換に基づく各データブロックのエレメントを抽出するステップ、をさらに含む請求項24に記載のコンピュータシステム。

【請求項26】 前記二面对称オペレーションが主対角

線上でデジタル画像をフリップすることを含む請求項24に記載のコンピュータシステム。

【請求項27】 前記二面对称オペレーションが中央垂直軸上でデジタル画像をフリップすることを含む請求項24に記載のコンピュータシステム。

【請求項28】 前記二面对称オペレーションが交差方向対角線上でデジタル画像を回転させることを含む請求項24に記載のコンピュータシステム。

【請求項29】 前記二面对称オペレーションが中央水平軸上でデジタル画像を回転させることを含む請求項24に記載のコンピュータシステム。

【請求項30】 前記二面对称オペレーションがデジタル画像を時計回り方向に90°回転させることを含む請求項24に記載のコンピュータシステム。

【請求項31】 前記二面对称オペレーションがデジタル画像を180°回転させることを含む請求項24に記載のコンピュータシステム。

【請求項32】 前記二面对称オペレーションがデジタル画像を反時計回り方向に90°回転させることを含む請求項24に記載のコンピュータシステム。

【請求項33】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内でエレメントを転置する請求項26に記載のコンピュータシステム。

【請求項34】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内で奇数列エレメントの符号を反転させる請求項27に記載のコンピュータシステム。

【請求項35】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内でエレメントを転置し、次いで各ブロック内のその他あらゆるエレメントの符号を反転させる請求項28に記載のコンピュータシステム。

【請求項36】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内で奇数行エレメントの符号を反転させる請求項29に記載のコンピュータシステム。

【請求項37】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内でエレメントを転置し、次いで各ブロック内の奇数列エレメントの符号を反転させる請求項30に記載のコンピュータシステム。

【請求項38】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内のその他あらゆるエレメントの符号を反転させる請求項31に記載のコンピュータシステム。

【請求項39】 前記一次変換ドメインオペレーションを前記複数の一次変換に基づく各データブロックに適用して、各ブロック内でエレメントを転置し、次いで各ブ

ロック内の奇数行エレメントの符号を反転させる請求項32に記載のコンピュータシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、一般には画像データの圧縮ドメイン表現の処理に関し、さらに詳しくは、画像データを完全な伸張および圧縮の工程に付さずに、画像の正規な幾何学的変換のような、ある空間ドメイン処理を達成するための、圧縮ドメイン表現の取り扱いに関する。

【0002】

【従来の技術及び発明が解決しようとする課題】典型的な高品質のデジタル化されたカラー画像では、RGBカラー空間における赤(R)、緑(G)および青(B)またはYCBCRカラー空間における輝度(Y)、クロミナンス(CB)およびクロミナンス(CR)のそれぞれに対して8ビットの、24ビット/画素(ビット・パー・ピクセル(bpp))が用いられている。このような画像を圧縮しない状態(例えば空間または画素ドメイン)で転送または格納することは、時間的および必要とするメモリの点から極めて費用が嵩む。そのため、高品質のデジタル化されたカラー画像を格納および/または転送するアプリケーションおよび装置、例えばデジタルカメラでは、現在入手可能な圧縮アルゴリズムを用いて、典型的には圧縮フォーマットで格納および/または転送が行われる。

【0003】JPEG(Joint Photographic Experts Groupの略語)のような圧縮基準の出現により、JPEG圧縮フォーマットのみで内容を生成、保存するデジタル・イメージング・システムおよびアプリケーションが標準である。例えば、Epson PhotoPC 600、Kodak DC-10などのようなほとんどのデジタル・スチル画像カメラ(DSCs)では、カメラが捕えた画像は直ちにカメラ内で圧縮され、そのカメラの格納システムにJPEGファイルとして格納される。しばしば、表示に先立って、これらの画像を操作する必要がある。典型的な画像の操作としては、(a)ポートレートから風景モードへの画像の回転またはその逆、(b)サイズの拡大または縮小のための画像のスケーリング、(c)画像における明度およびコントラストの変更、(d)新しい画像を作成する目的のためおよび合成操作するための画像の一部のクロッピング、(e)単純なビットマップ注釈の画像への付加、および(f)可視/不可視な透かしの画像への埋め込みが含まれる。デジタル・カメラの格納制限のために、これらの画像操作では処理される出力がJPEGフォーマットであることが要求される。

【0004】これらのタスク実行および圧縮モードでのみ画像を利用可能とするという必要から、圧縮ドメイン表現に直接応用可能な画像処理技術の開発に対する関心

が大いに高まっている。圧縮ドメイン処理法の研究は以下のような事実が端緒となっている。(a)圧縮ドメインにおけるデータの容量は空間ドメイン表現に比べて極めて小さいが、このことは、所望の処理タスクにとってサンプル当たりのオペレーションが少なく済むことを意味する。(b)データを伸張し、続いて空間ドメイン内で所望の画像処理関数を適用し、次いで転送または格納効率を高めるため再圧縮することが必要な慣用のパイプライン処理では、画像の忠実度の低下を招きかねない。さらにこのような慣用のパイプライン処理は、時として圧縮タスクが伸張タスクよりも複雑であるため、演算が極めて複雑化し、または待ち時間が長くなる。

【0005】これに反して、圧縮ドメインに基づく処理法では、ハフマン(Huffman)・デコーディングおよびハフマン・エンコーディングのような複雑さの低いタスクによりJPEG伸張および圧縮のタスクを置き換えるので、演算の複雑さが軽減されることがある

(S. F. ChangおよびD. G. Messerschmitt, 「Manipulation and Composition of MC-DCT Compressed Video (MC-DCT圧縮ビデオの操作および合成)」, IEEE JSAC Special Issue on Intelligent Signal Processing, vol. 13, 1~11ページ, 1995年1月; N. MerhavおよびV. Bhaskaran, 「A fast algorithm for DCT-domain inverse motion compensation (DCTドメイン逆転補償用の高速アルゴリズム)」, ICASSP '96, pp. IV. 2307~2310ページ, アトランタ, 1996年5月; B. NatarajanおよびV. Bhaskaran, 「A fast approximate algorithm for scaling down digital images in the DCT domain (DCTドメインにおけるデジタル画像縮小のための高速近似アルゴリズム)」, IEEE 画像処理国際会議(ICIP), ワシントンD. C., 1995年10月; およびBrian SmithおよびLarry Rowe, 「Algorithms for manipulating compressed images, (圧縮画像操作のためのアルゴリズム)」, IEEE Computer Graphics and Applications, pp. 34~42ページ, 1993年9月、参照)。

【0006】本発明の第1の目的は、空間ドメインにおけるデジタル画像の操作を行うのに伴う上述の問題を克服することにある。

【0007】本発明の第2の目的は、完全な伸張および圧縮の工程を経由すること無しに、画像の圧縮ドメイン

表現を直接操作して、選択された空間ドメイン操作を行うことにある。

【0008】本発明の第3の目的は、空間ドメインにおける対応する画像操作を行うのに必要な圧縮データ操作を大幅に単純化するアルゴリズムのセットを提供することにある。

【0009】本発明の第4の目的は、1セットの圧縮データを有し、かつ別のセットの圧縮データを生じて、工程逆転時には品質の低下を伴わずにオリジナル画像が生成される操作工程を含む、圧縮ドメイン内で画像データを操作するアルゴリズムのセットを提供することにある。

【0010】

【課題を解決するための手段】本発明では、デジタル画像の一次変換（線形変換）ドメイン表現を操作することにより、デジタル画像の空間ドメイン表現上で種々の二面角対称オペレーションを行う手法が提供される。デジタル画像がJPEGファイルのような圧縮ビットストリームの形式である場合、この手法には、圧縮ビットストリームをデコードして、デジタル画像の一次変換ドメイン表現を規定する、一次変換に基づくデータブロックの生成が含まれる。一次変換に基づくデータブロックは対応する幾何学的変換用に再配列され、特定の一次変換ドメインのオペレーションが各ブロック内のデータエレメントに適用され、そしてブロックが再アセンブリされる。これらのデータのブロックが空間ドメインに伸張されると、得られるデジタル画像はオリジナル画像に関してフリップまたは回転を生じる。

【0011】オリジナル画像については、生じた画像は対角線方向（主方向または交差方向のいずれか）に対するフリップ、中軸方向に対するフリップ（垂直または水平の何れか）、または90°、180°もしくは270°の回転を生じる。この手法はデジタルスチル画像カメラのような画像装置またはコンピュータシステムに適用することができる。いずれの場合も、ハードウェアまたはソフトウェアを用いてインプリメントすることができる。

【0012】本発明のその他の目的およびその詳細な内容については、添付の図面を参考として以下の説明およ

び特許請求の範囲から把握することができよう。

【0013】

【発明の実施の形態】JPEGのようなスチル画像圧縮基準における基底関数として8×8離散コサイン変換(DCT)が選ばれ(G. K. Wallace, 「The JPEG Still Picture Compression Standard」, Communication of the ACM, vol. 34, no. 4, 1991年4月を参照されたい)、またJPEGが市場の多くの一般向けデジタルスチルカメラに広く用いられている圧縮法であるため、以下、デジタルスチルカメラのアプリケーションに適し、かつカメラのJPEG圧縮ビットストリームのDCTドメイン表現上で動作する圧縮ドメイン処理法に絞って説明する。8×8が最も一般的なDCTブロックサイズであるが、これらの手法はその他のDCTブロックサイズにも拡張できる。さらに、これらの手法は、離散サイン変換、離散アダマール変換およびウェーブレット変換を含むその他の一次変換に基づく基底関数に拡張可能である。

【0014】JPEG圧縮および伸張工程ならびに画像処理関数をJPEGパイプラインに統合する考え方から簡単に説明する。JPEGではDCTを用いて、その空間または画素ドメイン表現からデータをより効率的にコード化し得る圧縮または周波数ドメイン表現にスチル画像データを変換する。ここで開発した画像操作法はDCTの特性の利点を活用するように設計されている。

【0015】JPEG圧縮および伸張工程は、図1に概略的に示す通り、それぞれ8×8のブロックサイズを有するブロック毎の基底上で動作する。図1にその概略を示す通り、非圧縮スチル画像11はラスタ・ブロック変換器12により8×8ブロックの画素に分解される。次いで、これらのブロックはフォワード8×8DCT13により転送されて、対応する8×8DCTブロックのセットを生成する。サンプル $f(i, j)$ の空間ドメイン8×8ブロックのフォワード8×8DCT、 $F(u, v)$ は以下のように算出される((1)～(3)式)。

【0016】

【数1】

$$F(u, v) = \frac{\alpha_u \alpha_v}{4} \sum_{i=0}^7 \sum_{j=0}^7 C_{i,u} C_{j,v} f(i, j), \quad (1)$$

ここで、

$$C_{x,y} = \cos \frac{(2x+1)y\pi}{16}, \quad (2)$$

$$\alpha_x = \begin{cases} 1, & x \neq 0 \\ 1/\sqrt{2}, & \text{その他の場合} \end{cases} \quad (3)$$

【0017】フォワード8×8DCTからの出力後、各64DCT係数は64エレメント量子化テーブルQに関連してフォワード量子化器14内で均一に量子化されるが、これは経験的に導出されて可視的には無意味の廃棄情報とされることができる。この圧縮工程で、圧縮の間に唯一生じるロス、 $F(u, v)$ から $F_Q(u, v) = \text{RoundOff}\{F(u, v)/Q(u, v)\}$ への量子化に起因する（ここで、Qは8×8量子化テーブルである）。

【0018】量子化の後、各ブロック内のDCTデータは、低周波係数（非ゼロとなる可能性大）を高周波係数（ゼロの可能性大）の前に持ってくることにより、エントロピー・コーディングを促進する「ジグザグ」順に並べ変えられる。次いで、データはハフマンエンコーダ15内でハフマンコード化して、データをさらにコンパクトにし、JPEG圧縮ビットストリームを生成するよう

$$f(i, j) = \sum_{u=0}^7 \sum_{v=0}^7 \frac{\alpha_u \alpha_v}{4} C_{i,u} C_{j,v} F(u, v), \quad (4)$$

【0021】圧縮工程は、量子化係数 F_Q により実際に働かせて f の近似値 f_Q のみを与える（（5）式）。

$$f_Q(i, j) = \sum_{u=0}^7 \sum_{v=0}^7 \frac{\alpha_u \alpha_v}{4} C_{i,u} C_{j,v} F_Q(u, v) F_Q(u, v), \quad (5)$$

【0023】回転またはスケーリングのような画像処理関数が、JPEG圧縮ビットストリームとしてしか得られない画像上で実行されねばならないとした場合には、この処理は図2に示すように行われる。先ず、JPEG圧縮ビットストリームがブロック31において、空間ドメイン表現に伸張し直される。次いで空間ドメイン処理がブロック32で実行される。その後、処理された画素データがブロック33で再圧縮されて新しいJPEG圧縮ビットストリームを生成する。この処理は伸張された空間ドメインデータ（画素）に適用されるので、この手法を空間ドメインアプローチと呼ぶことにする。

【0024】空間ドメインアプローチにはいくつかの利点がある。その1つは、空間ドメインでの画像処理はよく知られた問題であり、その解決法は多くの典型的な画像処理関数に広く用い得る点である。その他の利点は、処理関数がデータを表現するために用いる基本的な圧縮手法に依存しない点である。

にする。

【0019】画像は、対称的逆算法を用いて圧縮ビットストリームから再構築される。ハフマンデコーダ16内で圧縮ビットストリームをデコードすることによりJPEG圧縮工程が開始され、8×8ブロックのDCT係数が再び生成される。逆ジグザグ化の手続きで係数が再配列され、次いでブロックが逆量子化器17を通して与えられる。次の工程では、8×8逆離散コサイン変換器（IDCT）18がDCTの8×8ブロック上で動作して画素の8×8ブロックのストリームを生成する。ブロック・ラスタ変換器19がこれらのブロックを伸張スチル画像21に変換する。この伸張工程では、IDCTが係数 $F(u, v)$ を画素 $f(i, j)$ に正確に変換し直す（（4）式）。

【0020】

【数2】

【0022】

【数3】

【0025】このアプローチには欠点もある。画像処理関数を適用する前に、データは完全に伸張されていなければならない。さらに、処理されたデータは再度圧縮工程に付す必要がある。JPEGはロスの多い圧縮法であるので、伸張一再圧縮により画像の品質低下を招く恐れがある。ポートレートから景観モードへの切り換えアプリケーションについては、向きが変更される度に画像品質は低下の一途をたどる。その他の欠点は、伸張および圧縮の複雑さが極めて著しいことである。例えば、図2で画像処理タスクが仮に時計回り方向に90°回転とすると、入力データの各8×8ブロックは表1のようになる。オペレーションカウンタの大まかな推計を行うため、積算、加算およびデータアクセスがそれぞれ1データ項目当たり1オペレーションで実行されるものと仮定した。

【0026】

【表1】

タスク	オペレーション カウント	コメント
ハフマンデコード	81	IDCTの複雑さを15%と推定
逆量子化	64	1積算に1オペレーションを要すると推定
8×8 DCT	544	80積算し464を加算（[*]による）
90°の回転	128	単純行列転置 （1画素当たり1リード、1ライト）
8×8 DCT	544	80積算し464を加算（[*]による）
フォワード量子化	64	1積算に1オペレーションを要すると推定
ハフマンエンコード	54	IDCTの複雑さを10%と推定
トータル	1415	

表1：入力および出力データが8×8 DCT形であるときの空間ドメインに基づく90°回転オペレーションのオペレーションカウント。表中[*]印はY、アライ、T、アグイおよびM、ナカジマによる「画像のための高速DCT-SQ手法」、IEICE会報、E71(11)：1095、1988年11月から引用。

【0027】本発明は空間ドメインに基づく画像処理に代わる、すなわち圧縮ドメインに基づく画像処理を提案するものである。後者は、すでに圧縮形で得られているデータ、例えばJPEGビットストリームによく適合する。圧縮ドメインに基づく画像処理手法の基本的処理は図3に示す通りである。JPEGデータについては、圧縮ドメインに基づく画像処理として、通常、図3においてブロック42で示すDCTドメイン画像処理が含まれる。

【0028】図3に示す通り、DCTドメイン画像処理42にはブロック41におけるJPEG圧縮ビットストリームのエントロピーデコーディングが先行し、そしてブロック43におけるエントロピーコーディングが後続する。

【0029】同じく図3に示す通り、DCTドメイン画像処理ブロック42はさらに44～49を付番された処理ブロックに分割される。前述の通り、ビットストリームの伸張によるDCT係数の8×8ブロックが得られる（44）。これらの係数は、ジグザグ解除の手続き（45）を用いて再配列され、その後、DCT係数のブロックが量子化解除される（46）。本発明によれば、DCT係数の量子化解除ブロックはブロック指向処理を行う（47）。この処理に続いてDCTデータに基づいてブロックが量子化され（48）、そしてジグザグ順に再配列される（49）。しかし、画像処理タスクに依存してすべてのブロック44～49について行われる必要はない。例えば、D₄オペレーション（対角線およびY軸に

関するフリップの順で表わすことができる）については、ジグザグ解除45、量子化解除46、ジグザグ化49、および量子化48の各工程が省略可能である。圧縮ドメインに基づくD₄オペレーションの詳細については後述する。

【0030】一般に、圧縮ドメイン処理には以下のような利点がある。まず、多くの場合、量子化解除—量子化工程が省略できるので、画像品質を保つことができる。第2に、完全なJPEG伸張および圧縮のタスクがなくなるので、図2に示す空間ドメインの場合よりも複雑さが大幅に低い。特に90°回転の場合は、表1の圧縮ドメインに基づく処理相手先のインプリメントは表2に示すオペレーションカウントを有する。圧縮ドメインアプローチを用いる全オペレーションカウントは表1の空間ドメインに基づくアプローチのほぼ1/5以下であることに注目すべきである（DCTドメインに基づく回転法については後述する）。その他の利点としては、典型的なデジタル画像表現では、画素の深い相互関係のためにDCTドメイン表現が極めて散在（例えば8×8 DCTブロックで通常7～16の非ゼロ値）している点が挙げられる。このデータの散在性は、DCTドメイン処理アプローチの活用により、全体的な複雑さをさらに軽減することができるが、この特性は空間ドメイン表現では得られない（D₄オペレーションに対するこの特性の活用については後述する）。

【0031】

【表2】

タスク	オペレーション カウント	コメント
ハフマンデコード	81	IDCTの複雑さを15%と推定
逆量子化	—	不要
8×8 DCT	—	不要
90° の回転	32	全奇数列の符号変更を要する
	112	行-列データ交換を要する
8×8 DCT	—	不要
フォワード量子化	—	不要
ハフマンエンコード	54	IDCTの複雑さを10%と推定
トータル	279	

表2：入力および出力データが8×8 DCT形のときの圧縮ドメインに基づく90° 回転オペレーションのオペレーションカウント。

【0032】一般に、空間ドメインに基づく画像処理関数に匹敵する、圧縮ドメインに基づく処理を導出することは不可能である点に注意すべきである。DCTは一次変換であり、従って圧縮ドメインに基づく処理は一次画像処理関数について行われると予想される。メディアンフィルタリング、変形/異形化のような非線形画像処理関数は、図3に示すもののように圧縮ドメインに基づくアプローチには適合しない。

【0033】《JPEG画像上での二面对称オペレーション》以下、JPEG圧縮データの簡単な幾何学的変換

を制御する基本方程式について説明する。対角軸、Y軸（すなわち、中央垂直軸）およびX軸（すなわち、中央水平軸）に関するフリップの構成で規定されるオペレーションにより、 D_4 と呼ばれる四辺形の二面对称のグループが形成される。これらのオペレーションについて表3に整理して説明する。なお、表3および表4において、および以下の説明において、オペレーションを説明の便宜上〈 〉を用いて示す。

【0034】

【表3】

オペレーション	説明	〈 F_d 〉および〈 F_y 〉による生成
〈I〉	本体	〈 F_d 〉 〈 F_d 〉
〈 F_d 〉	対角線フリップ（主対角線についてのフリップ）	〈 F_d 〉
〈 F_y 〉	列フリップ（y軸についてのフリップ）	〈 F_y 〉
〈 F_{cd} 〉	交差対角線フリップ（交差対角線についてのフリップ）	〈 F_y 〉 〈 F_d 〉 〈 F_y 〉
〈 F_x 〉	行フリップ（x軸についてのフリップ）	〈 F_d 〉 〈 F_y 〉 〈 F_d 〉
〈 R_{90} 〉	時計回り90° 回転	〈 F_y 〉 〈 F_d 〉
〈 R_{180} 〉	180° 回転	〈 F_y 〉 〈 F_d 〉 〈 F_y 〉 〈 F_d 〉
〈 R_{-90} 〉	反時計回り90° 回転	〈 F_d 〉 〈 F_y 〉

表3：四辺形の二面对称の D_4 グループ

【0035】全体のグループを生成するようにオペレーション〈 F_d 〉および〈 F_y 〉を構成することができる点に注意すべきである（注：オペレーション〈 o_1 〉および〈 o_2 〉の構成〈 o_1 〉〈 o_2 〉は、先ず〈 o_2 〉を、次いで〈 o_1 〉を適用して得られるオペレーションである）。例えば、単純に時計回り方向に90° 回転（オペレーション〈 R_{90} 〉）は、図4に示す通り、対

角線フリップを適用し、次いで列フリップを適用することにより達成できる。従って、二つのオペレーション〈 F_d 〉および〈 F_y 〉の圧縮ドメインに基づく相手先が導出できれば、残りもすべて導出可能である。

【0036】 f を8×8画素ブロックとし、 F を対応する8×8 DCTブロックとする（ $DCT(f) = F$ および $IDCT(F) = f$ ）。8個の D_4 オペレーション

〈o〉について、fと〈o〉fの関係を表わすのは容易である。目標はFとDCT (〈o〉f) の関係 (これを〈o〉Fと表わす) の導出である。

【0037】オペレーション〈F_y〉の場合、空間ドメインでは列フリップ出力〈F_y〉f(i, j)が以下のように表わされる ((6)式)。

【0038】

【数4】

$$\langle F_y \rangle f(i, j) = f(i, 7-j) \quad (6)$$

【0039】また、(1)式から〈F_y〉F(u, v)は、以下のように表される ((7)~(9)式)。

【0040】

【数5】

$$\langle F_y \rangle F(u, v) = \frac{\alpha_u \alpha_v}{4} \sum_{i=0}^7 \sum_{j=0}^7 C_{i,u} C_{j,v} \langle F_y \rangle f(i, j) \quad (7)$$

$$= \frac{\alpha_u \alpha_v}{4} \sum_{i=0}^7 \sum_{j=0}^7 C_{i,u} C_{j,v} f(i, 7-j) \quad (8)$$

$$= \frac{\alpha_u \alpha_v}{4} \sum_{i=0}^7 \sum_{k=0}^7 C_{i,u} C_{7-k,v} f(i, k) \quad (9)$$

(ここで、k=7-j)

【0041】式(2)から、C_{7-k, v}は、次のように表される ((10)~(14)式)。

【0042】

【数6】

$$C_{x,y} = \cos\left(\frac{\{2(7-k)+1\}v\pi}{16}\right), \quad (10)$$

$$= \cos\left(v\pi - \frac{(2k+1)v\pi}{16}\right) \quad (11)$$

$$= \cos(v\pi) \cos\left(\frac{(2k+1)v\pi}{16}\right) \quad (12)$$

$$= (-1)^v \cos\left(\frac{(2k+1)v\pi}{16}\right) \quad (13)$$

$$= (-1)^v C_{k,v} \quad (14)$$

【0043】(14)式、(1)式のDCT定義から、(7)~(9)式における、〈F_y〉F(u, v)は、以下のように書き換えることができる ((15)、(16)式)。

【0044】

【数7】

$$\langle F_y \rangle F(u, v) = \frac{\alpha_u \alpha_v}{4} (-1)^v \sum_{i=0}^7 \sum_{k=0}^7 C_{i,u} C_{k,v} f(i, k), \quad (15)$$

$$= (-1)^v F(u, v) \quad (16)$$

【0045】空間ドメインでは、入力ブロックの対角線フリップf(i, j)は〈F_d〉f(i, j)=f(j, i)となる。このDCTドメインの式は以下のように表わされることは明らかである ((17)式)。

【0046】

【数8】

$$\langle F_d \rangle F(u, v) = F(v, u) \quad (17)$$

【0047】式15および式17、ならびに表3の第3

列で与えられる関係を用いて、 D_4 オペレーションのすべてについて圧縮ドメインの相手先を導出することができる。これらを表4に示す。

【0048】

【表4】

空間ドメイン	DCTドメイン	入力ブロック処理
$\langle I \rangle f(i, j) = f(i, j)$	$\langle I \rangle F(u, v) = F(u, v)$	なし
$\langle F_d \rangle f(i, j) = f(j, i)$	$\langle F_d \rangle F(u, v) = F(v, u)$	転置
$\langle F_y \rangle f(i, j) = f(i, 7-j)$	$\langle F_y \rangle F(u, v) = (-1)^v F(v, u)$	奇数列の符号反転
$\langle F_x \rangle f(i, j) = f(7-j, 7-i)$	$\langle F_{cd} \rangle F(u, v) = \langle F_y \rangle \langle F_d \rangle \langle F_y \rangle F(u, v)$ $= (-1)^v \langle F_d \rangle \langle F_y \rangle F(u, v)$ $= (-1)^v \langle F_y \rangle F(v, u)$ $= (-1)^v (-1)^u F(v, u)$	その他すべての エレメントの転置 および符号反転
$\langle F_x \rangle f(i, j) = f(7-i, j)$	$\langle F_x \rangle F(u, v) = \langle F_d \rangle \langle F_y \rangle \langle F_d \rangle F(u, v)$ $= \langle F_y \rangle \langle F_d \rangle F(v, u)$ $= (-1)^u \langle F_d \rangle F(v, u)$ $= (-1)^u F(u, v)$	奇数行の符号反転
$\langle R_{90} \rangle f(i, j) = f(7-j, i)$	$\langle R_{90} \rangle F(u, v) = \langle F_y \rangle \langle F_d \rangle F(u, v)$ $= (-1)^v \langle F_d \rangle F(u, v)$ $= (-1)^v F(v, u)$	奇数列の転置 および符号反転
$\langle R_{180} \rangle f(i, j) = f(7-i, 7-j)$	$\langle R_{180} \rangle F(u, v) = \langle F_y \rangle \langle F_d \rangle \langle F_y \rangle \langle F_d \rangle F(u, v)$ $= (-1)^v \langle F_d \rangle \langle F_y \rangle \langle F_d \rangle F(u, v)$ $= (-1)^v \langle F_y \rangle \langle F_d \rangle F(v, u)$ $= (-1)^v (-1)^u \langle F_d \rangle F(v, u)$ $= (-1)^v (-1)^u F(u, v)$	その他すべての エレメントの符号 反転
$\langle R_{-90} \rangle f(i, j) = f(j, 7-i)$	$\langle R_{-90} \rangle F(u, v) = \langle F_d \rangle \langle F_y \rangle F(u, v)$ $= \langle F_y \rangle F(v, u)$ $= (-1)^u F(v, u)$	奇数行の転置 および符号反転

表4: D_4 における幾何学的変換のDCTドメインオペレーション

【0049】上記の関係は量子化を考慮せずに導出したものである。 $F(u, v)$ の代わりに単純に量子化係数 $F_Q(u, v)$ $Q(u, v)$ を用いることにより、以下のことが明らかになる。すなわち、 D_4 オペレーションで生成するブロックの量子化係数は、対応する入力ブロック処理(表4の3列)を元のブロックの量子化係数に適用することによって直接得ることができる。量子化テーブルは同一のままである($\langle F_d \rangle$, $\langle F_{cd} \rangle$, $\langle R_{90} \rangle$ および $\langle R_{-90} \rangle$ について変換される)。このDCTドメインアプローチによりIDCT、DCT、ならびに量子化解除および量子化が回避される。ジグザグ化解除およびジグザグ化工程も回避されることについては後述する。

【0050】多くの 8×8 ブロックからなる $W \times H$ のJPEG画像上で、 D_4 オペレーションを行うには、圧縮ドメインに基づく手続きは以下の通りである。(a)対応する幾何学的変換について 8×8 DCTブロックの再配列、および(b)各 8×8 DCTブロック内のエレメントに対して表4によるDCTドメインに基づくオペレーションを適用することである(なお、JPEG画像における W および H は8の倍数であり、必要に応じてオリ

ジナル画像のパディングによって得られる)。(a)および(b)はロスのないオペレーションであって、量子化DCT係数 $F_Q(u, v)$ が符号変更を越えて処理されないことに注意すべきである。すなわち、量子化解除および再量子化の工程を回避することにより、カメラのJPEGファイルに対して表4の1種または2種以上の幾何学的変換を何回も適用しても、品質のロスを生じることはない。

【0051】単純化するため、グレースケール画像を用いるインプリメント法について説明し、続いてカラー画像に必要な簡単な修正を概説する。ここで、画像の幅および高さを8の倍数とする。さらに、カラー画像については、サブサンプリングはブロックパディングを必要としないものとする。これらの制約が満足されない場合、ある別の行/列のオリジナル画像を改めてパディングすることにより、ここに述べたすべてのオペレーションを適用することができる。

【0052】JPEGデータとして得られる $W \times H$ グレースケール画像 I について考察する。JPEGデータにはエントロピー・コーディングを適用し、DC項には差分コーディングを取消すことにより、各ブロックについ

ての量子化DCT係数を得ることができる。 F_k は画像についての量子化DCT係数の 8×8 ブロック(ラスタ順の付番 k)を表わすものとする($0 \leq k < WH/64$)。

【0053】 I_0 は画像にオペレーション〈 o 〉を適用した結果を表わすものとする(ここに、〈 o 〉は表3からの D_4 オペレーションの1つである)。以上のことは、 I_0 のDCT係数のブロックは、可能な再配列転置および符号変更を有する I と本質的に同じであり、そして量子化テーブルも同じであり、可能な転置を有することを示している。一般に、 I_0 における量子化係数のブロック F_{k_0} は次のように表わされる。

【0054】 $F_{k_0} = \langle o \rangle F_p \langle o \rangle (k)$

ここに、 $p \langle o \rangle$ はブロックの順列である。

【0055】一例として時計回り方向に 90° 回転(〈 o 〉= R_{90})について考察する。 $k = i_b(H/8) + j_b$ (回転画像 I_0 における i_b 行および j_b 列のブロック)については、

$p \langle o \rangle (k) = (H - j_b - 1)(W/8) + j_b$ となる。

【0056】JPEG画像 I_0 を形成させるには、 $k = 0, 1, 2, \dots$ 順のブロック F_{k_0} を計算する必要がある。この計算では $p \langle o \rangle (0), p \langle o \rangle (1), p \langle o \rangle (2), \dots$ 順(これは、一般に、ブロックがJPEG画像 I に格納されるラスタ順とは異なる)のブロック I をアクセスする必要がある。圧縮ブロックはサイズが任意に異なるため、JPEGビットストリームからの特定のブロックを抽出するにはその前のブロックをすべて解析しなければならず、極めて費用が嵩む。さらに、量子化された特異的コードのために、前のDC値がすべてデコードされた後、ブロックのDC係数の値を初めて抽出することができる。二つのパスアプローチを用いて、この問題を回避する。

【0057】最初のパスでは、JPEG画像 I を解析して各ブロックのビット・オフセットおよび各ブロックのDC値を抽出する。第2のパスでは、各画像 I_0 について F_{k_0} を演算する際、ブロック数についてのビット・オフセットをルックアップし、かつ I についてのJPEGビットストリームの位置をシークすることにより、 $F_p \langle o \rangle (k)$ がすぐにアクセスされる。このアルゴリズムは疑似コードで次のようにまとめることができる。

【0058】procedure Operate
input: JPEGImage I, オペレーション

〈 o 〉

output: JPEGImage I_0

/*第1パス: オフセットおよびDC値を集める*/
 I からのヘッダーを解析する。

for $k=0$ to $WH/64-1$

1. オフセットをブロック k のビット・オフセットとして I に格納する。2. DC項を抽出しつつ I を解析して過去のその他のブロックを移動する。3. ブロック k について特異的コード化を取消しおよびDC値の格納する。

【0059】/*第2パス: I_0 の演算および格納*/
 I_0 のヘッダーの演算・格納を行い、〈 o 〉により要求されるなら I の量子化テーブルを転置する。

for $k=0$ to $WH/64-1$

1. I をシークし、ブロック数 $p \langle o \rangle (k)$ の量子化係数を抽出するため、オフセットおよびDCテーブルを使用する。2. 必要に応じて、 o による係数への符号変更の転置および/または適用する(表4の第3列)。3. $I \langle o \rangle$ についてのJPEGビットストリームへのエントロピー・コード化係数を貼り付ける。

【0060】《ブロックの効率的な転置および符号変更》元のブロック $F_p \langle o \rangle (k)$ が o 経由で転置されて画像 I_0 についてのブロック F_{k_0} を得る際、「オペレートJPEG」アルゴリズムの心臓部は第2パスにある。 8×8 ブロックに符号変更の転置および適用を行うには、ストレート・フォワード方式でインプリメントされていれば、ブロックの各エレメントを一度アクセスする必要がある。つまり、複雑さは64の倍数となる。典型的なJPEG画像における量子化係数ブロックは極めて少ない数(典型的には16以下)の非ゼロ係数を有する。ここに、複雑さがブロック内の非ゼロ係数の数に比例する、転置および符号変更を行うためのアルゴリズムを提示する。このアルゴリズムにより、全体の複雑さが実質的に減り、圧縮画像の全サイズでは非圧縮画像の全サイズにおけるよりも線状になる。

【0061】JPEGのハフマンコード化モードでは、量子化係数のブロックはジグザグ順にスキャンすることによりコード化されて、ゼロの長いランを一まとめにグループ化する。このジグザグ配列を表5に示す。ブロックをコード化するとき、これらの差分に固有のハフマンテーブルを用いて、先ず現在のブロックと前のブロックの量子化DC値との差がコード化される。

【0062】

【表5】

	0	v →							7
0	0	1	5	6	14	15	27	28	
	2	4	7	13	16	26	29	42	
	3	8	12	17	25	30	41	43	
u ↓	9	11	18	24	31	40	44	53	
	10	19	23	32	39	45	52	54	
	20	22	33	38	46	51	55	60	
	21	34	37	47	50	56	59	61	
7	35	36	48	49	57	58	62	63	

J P E Gで用いられるZ Z (u, v)ジグザグ配列

【0063】次いで、量子化AC係数がジグザグ順にスキャンされ、そして(R, S, V)形式の連続する三つの記号としてコード化される。ここに、Rはジグザグ配列の連続するゼロの数であり、そして次の非ゼロ係数は以下のような値xを有する。

【0064】

【数9】

$$S = \lceil \log_2(|x|+1) \rceil$$

$$V = S \left\{ \begin{array}{l} x > 0 \text{ のとき, } x \\ x < 0 \text{ のとき, } x - 1 \end{array} \right\} \text{ の LSB }.$$

【0065】xまたはx-1の標準相補2進表現(standard 2's complement binary representation)を用いてVのビットが抽出される。J P E Gビットストリームでは、各(R, S, V)が、先ず(R, S)をエンコードするハフマンテーブルを用い、次いでVをエンコードするSエキストラ・ビットを用いてエンコードする。ゼロおよび終端ブロックの極めて長いプログラム実行(run)のような、いくつかの特別なエンコード条件では、特殊なコードでの処理が行われる。しかし、ここでの目的にとって、J P E Gデータは容易に解析されて次のようなデータ構造で各ブロックを捕捉することは明らかである。

【0066】

```
struct Jpeg Block
{
    int D; /*量子化DC係数*/
    int N; /*非ゼロAC係数の数*/
    struct {
        int Z; /*ジグザグインデックス*/
        int S;
        int V;
    } A[63];
};
```

このデータ構造中、Nはブロック内の非ゼロAC係数の数である。アレイAの最初のNエントリーのみ意味がある。エレメントA[k]はジグザグ位置(Z)を与え、そしてSおよびVはk番目の非ゼロAC係数 $0 \leq k < N$ を評価する。J P E Gビットストリームからこの構造へのブロックの読取り、およびJ P E Gデータとしてのこの構造の書込みは、いずれもストレートフォワードであり、ここでは詳細については省略する。

【0067】オペレーション〈o〉では転置および/または符号変更が必要となる。上記のデータ構造を用いることにより、〈o〉をインプリメントしてNの工程のみを(64よりも)必要とする新規なアルゴリズムについて説明する。このアルゴリズムはジグザグ順のブロック表現を直接用いるので、係数のジグザグ解除およびジグザグ化が回避される。

【0068】入力「JpegBlock」構造B(ブロックFに対応)を与えて、〈o〉Fに対応する構造B0を求める。非ゼロAC係数xの符号を変更するには、対応するSおよびV値は S_x および V_x であり、単純に V_x のビット補数を用いれば十分である。すなわち、

$$S - x = S_x, \quad V - x = -V_x$$

s〈o〉[64]をブール・フラグの予め演算されたアレイとして、もし〈o〉が、変更されるべきZ番目のジグザグ係数の符号を必要とすれば、s〈o〉[Z]がTRUEとなるとする。符号の変更のみが必要とされれば(すなわち、オペレーション〈F_x〉、〈F_y〉、〈R

180) についての変更)、s (0) [A [k]. Z] がTRUEであるA [k]. Vについてビットをフリップ化させつつ、各エントリーをコピーすることにより、BをB (0) に変換することができる。

【0069】N工程で「JpegBlock」構造の転置を行うには、以下のキー観測を用いる。ある交差方向対角線上でのエレメントが正順に反転されている場合を除き、アレイAは同じままである。これを説明するため、t [Z] が元のジグザグ・インデックスZに対する転置ジグザグ・インデックスを表わすとする。すなわち、

t [1] = 2, t [2] = 1, t [3] = 5, t [4] = 4, ...

ジグザグ係数が付番された場合を考えると、

1, 3, 6, 7, 10, 11, 14

が唯一の非ゼロ係数である。従って、転置後のジグザグ配置は次のようになる。

【0070】

```

procedure Transpose And SignChange
input: JpegBlock B, arrays t, s (0), d
output: JpegBlock B (0)
B (0). D = B. D /* s (0) [0] は常にFALSEであることに注意 */
B (0). N = B. N
if (B. N == 0) then return
dcur = d [B. A [0]. Z] /* 開始時の交差対角線数 */
startpos = 0 /* [B. A [0]. Z] /* 交差対角線数 dcur
が B. A [...] で開始する位置 */
for k = 0 to B. N
    /* 新しい交差対角線数を見出す */
    if (k < B. N) then dnew = [B. A [k]. Z]
    else dnew = ∞ /* 14より大きい値は、まさに dnew > dcur
    であることを確認 */
    if (dnew < dcur) then
        /* B. A [startpos... (k-1)] を B (0). A [...]
        に逆順にコピー */
        for l = 0 to k - startpos - 1
            i = k - 1 - l /* B. A [...] におけるインデックス */
            j = startpos + l /* B (0). A [...] におけるインデ
            ックス */
            B (0). A [j]. Z = t [B. A [i]. Z]
            B (0). A [j]. S = B. A [i]. S
            B (0). A [j]. V = B. A [i]. V
            if (s (0) [B (0). A [j]. Z]) then
                B (0). A [j]. V = ¬ B (0). A [j]. V /* ビッ
                トを反転 */
        /* dcur と startpos を更新 */
        dcur = dnew
        startpos = k
《カラー画像》

```

【数10】

t[1], t[3], t[7], t[6], t[14], t[11], t[10].

【0071】従って、同じ交差方向対角線上にある係数の各グループは反転させる必要がある。これを、非ゼロ係数のアレイB. A [...] をスキャンし、交差方向対角線が変化する各係数を解析することにより、効率的に行って、B. A [...] からB (0).

A [...] に逆順に前の交差方向対角線をコピーする。

【0072】以下の疑似コードがこのアルゴリズムを要約するものである。アレイtおよび(上述の)s (0)は予め演算される。さらに、d [64]を別の予め演算されたアレイとすると、各ジグザグ・インデックスについて「交差方向対角線の数」が得られる。つまり、ジグザグ・インデックスZが行数uおよび列数vに対応すると、d [Z] = u + vとなる。

【0073】

一般に、J P E G画像Iは1つ以上のカラー平面からなり、ある平面はサブサンプル化されている。カラー平面の数をPで表わすとする。各平面p（ただし、 $1 \leq p \leq P$ ）に伴って、水平サンプリング要素 w_p および垂直サンプリング要素 h_p がある。平面数pの幅および高さはそれぞれ Ww_p/w_{max} および Hh_p/h_{max} で与えられる。ここに、全平面（ $1 \leq p \leq P$ ）にわたって、 w_{max} が w_p の最大値であり、 h_{max} が h_p の最大値である。

【0074】J P E Gデータはさらにスキャンからなる層を有することができる。1スキャンは1または2以上のカラー平面からなり、制限されたビット精度の係数値を有する。「Operate J P E G」アルゴリズムは、単純に各スキャンに連続して適用することにより、拡張することができる。

【0075】各スキャン内では、スキャンにおける全カラー平面のブロックからのデータを、最小コード単位(MCU)として知られる単位にまとめられる。各MCUは、すべての w_p および h_p で決定される、固定順に並んだ固定数のブロックからなる。ここで取り扱うオペレーションについては、サンプリング要素を変更しない限り、画像I⁰はIとまったく同じスキャンおよびMCUのグループ分けである(w_p および h_p は転置を要するオペレーションに対してスワップされる必要がある)。MCUの順および各MCU内のブロックの順は変動することがあり得るが、I⁰における各MCUはIにおける正に1つのMCUから得られる。この事実を活用して、ビット・オフセットおよび、ブロックレベルよりもMCUレベルのDCテーブルを設けることにより、メモリの節減が図れる。

【0076】このように、Iにおける各スキャンについては「Operate J P E G」が2つのパスで動作する。最初のパスでは、そのスキャンの各MCUについてビット・オフセットが再コード化され、そして当該MCUにおける各カラー平面からの最初のブロックのDC値が再コード化される。第2のパスでは、I⁰のスキャンの特定のMCUが演算されるとき、対応するIのMCUが抽出され、そして各構成要素ブロックの「Jpeg Block」構造が完成する。次いで、これらのブロックは(オペレーションの要求に応じて)転置/符号変更/再コード化されて、J P E Gデータで書かれたI⁰のためのMCUを生成する。

【0077】《圧縮ドメインに基づく画像回転-DSCアプリケーション》デジタルスチル画像カメラ(DSC)を含む各種デジタル装置との関連して、本発明の圧縮ドメイン処理法が用いられる。DSCのブロック図を図5に示す。DSC60はマイクロプロセッサの制御下に操作されるが、画像を捕捉し、そしてブロック61においてアナログ電気信号に変換する電荷結合デバイス(CCD)を有する。次いで、このアナログ信号は、ブ

ロック62で処理され、デジタル化画像が一時的にフレームバッファ63に格納され、一方、ブロック64でデジタル処理に付される。デジタル画像処理ブロック64は、圧縮および伸張を含むいくつかの機能を行い、また本発明の圧縮ドメインに基づく処理法も実行する。処理ブロック64は、ユーザーコントロール65の下に、伸張された画像データが格納されるカメラ内画像ストレージ66とインターフェースを有する。ストレージ・ブロック66は、いずれもDSC60内で取り外し可能または固定型のコンパクト磁気媒体または固体媒体を内蔵し、そして取り外し可能な大容量PCMCIAフォーマット・ハードディスクカードまたはフラッシュ・メモ리카ードを用いることができる。

【0078】DSC60はアナログ出力器67およびデジタル出力器68を内蔵しており、これらを通して画像データがDSC内または外部装置に転送される。圧縮されていない画像データは、アナログ出力器67を経由して、DSC60内のLCDスクリーン69に転送されるか、またはVCRやTVモニターのような外部装置に転送される。また、圧縮または非圧縮画像データは、デジタル出力器68を通して、画像を表示するコンピュータシステムのようなデジタル装置にも転送される。

【0079】DCTドメイン表現を直接操作して90°、180°、270°回転またはミラーフリップのようなD₄オペレーションを実行する能力はDSC60では極めて有用である。カメラを垂直に保持して背の高い被写体を捕捉する場合、得られる画像は、水平にしたカメラで撮影した同じ写真を90°回転させた画像として現われる。典型的なフィルムカメラでは、このことは問題にならない。何故なら、正しい画像を見るには、写真を90°回転させるだけで済むからである。上述のようにDSC60で撮影した写真は、直接画像が焼付けられるか、またはTVモニター上に写されたり、コンピュータドキュメントに組み込まれるため、問題を生じる恐れがある。

【0080】DSC画像ファイルの読み出し中に、圧縮ドメインに基づく画像処理を組み込むことにより、ここに記載する方法で回転を取り消してDSC60が修正画像を生じるようにすることができる。DSCシステムでは、回転機能はDSC内にあってもよく、またはDSC60に接続されたコンピュータ上で動作するソフトウェア内にあってもよい。制限された演算能力および制限されたメモリを有するDSC60については、圧縮ドメインに基づくアプローチが、幾何学的転送を行うための唯一の効率的手段である。データを伸張し、次いで操作を行う別のアプローチはメモリ集約型であり、DSC60内での遅いCPU速度のために著しい能力低下を来す恐れがある。DSCシステムの進展に伴い、画像強化および透かし挿入のような付随的特性が、圧縮ドメイン処理

の枠内に容易に組み込まれるようになり、しかも圧縮ドメイン表現と空間ドメイン表現の間を行き来するには及ばない。

【0081】上述の通り、図5に示すDSC60は、コンピュータシステム、その他、デジタル画像を捕捉、処理及び検証する構成要素との関連でも用いることができる。図6は、DSC60、コンピュータシステムおよびその他の各種構成要素の間の相互関係を説明するブロック図である。参照番号100で一般的に表記するコンピュータシステムは、メインフレームまたはパーソナルコンピュータのようないかなる適当な形式のものであってもよい。

【0082】コンピュータシステム100は、慣用のマイクロプロセッサである中央処理装置（CPU）101、情報を一時的に蓄積するためのランダムアクセスメモリ（RAM）102、および情報を永久蓄積するためのリードオンリーメモリ（ROM）103を含む。これらの各構成要素はバス104と結合されている。コンピュータシステム100のオペレーションは典型的にはオペレーティングシステムソフトウェアによって制御、調整される。システムメモリに組み込まれ、CPU101上で動作するオペレーティングシステムは、システムリソースの配置を制御し、種々のタスク、例えば、特に処理、メモリ管理、ネットワークおよびI/O機能、その他を行うことにより、コンピュータシステム100のオペレーションを調整する。

【0083】また、ディスク107のような不揮発性大量記憶装置を挿入するディスク・ドライブ106もコントローラ105によってバス104に接続されている。同様に、コントローラ108は、バス104とCDROM110挿入用のコンパクトディスク（CD）ROMドライブ109との間にインターフェースを有する。ハードディスク111は、ディスクコントローラ113によりバス104に接続された固定ディスクドライブ112の一部として設けられている。

【0084】圧縮ドメインに基づく処理法のためのソフトウェアは記憶装置107および110上に格納され、そしてCPU101に転送、実行される。これとは別に、ソフトウェアをRAM102またはROM103に格納してもよい。同様に、本発明によりすでに処理され、または処理されるべき画像データは、ディスク107およびCDROM110のような取り外し可能な記憶媒体を用いて、コンピュータシステム100にロードし、またはこのシステムから抽出することができる。

【0085】画像データは別の手段でコンピュータシステム100入力してもよい。フィルムカメラ115で生成されたフィルムベースの画像114は、格納のために、スキャナ116でデジタル化し、そしてコンピュータ100で処理することができる。前述の通り、DSC60が画像を直接デジタル化し、コンピュータ100

0に転送してもよい。コントローラ123を経由してバス104に接続するキーボード121およびマウス122は、このようなデータの入力を容易にする。さもないと、コンピュータシステム100に情報を入力するための手段を設けることになる。

【0086】また、画像データは遠隔地向けにコンピュータ100へ、またはコンピュータ100から転送することができる。その目的のために、コンピュータ100は、直接接続またはモデム経由でネットワーク125との通信を可能とする通信アダプター124を含むことができる。なお、ネットワークにはローカルエリアネットワーク（LAN）、インターネットまたはオンラインサービスが含まれる。

【0087】コンピュータ100に転送または格納されたデジタル画像は多種多様な方法で見ることができる。コンピュータ100に付属するプリンタ126によりカラープリントが得られるが、その画質はプリンタ126によって変化がある。その他のオプションとしては、コンピュータ100に付属するディスプレイ127上で画像を見ることである。さらに別の選択としては、vcrを用いてテレビ受像器上に画像を表示することができる。

【0088】上述の通り、JPEG圧縮スチル結像のための簡易な圧縮ドメインに基づく処理フレームワークが本発明者等により開発された。鏡像、90°、180°および270°回転像のような簡易な幾何学的変換が、画像の忠実度を失わずにDCTドメイン内で容易に行い得ることが明らかになった。さらに、90°回転のような単純な変換であっても、慣用の空間ドメインに基づく処理アプローチに代わるDCTドメインで行った場合、著しく迅速に達成できることも明らかになった。現実的な設定における圧縮ドメインに基づく画像変換の実用例として、特にデジタルスチルカメラについても説明した。デジタルカメラが捕捉する画像は初めはノイズが極めて多いので、これらの画像に適用されるいかなる画像処理もノイズを増加させないことが命題であるが、ここに開発された方法は本質的に画像ロスなしのオペレーションである。

【0089】さらに、これらの方法は、一般向けデジタルスチルカメラに見られる、制限された演算およびメモリ能力に十分適合するものである。この基本的な圧縮ドメインフレームワークはその他の形式の画像変換にも用いることができ、殊にここに開発された方法論は、画像強化、画像ろ過に拡張できる（B. Chitpraser および K. R. Rao）、「離散コサイン変換ろ過」、信号処理、19巻、233-245ページ、1990年参照）。

【0090】特定の具体例との関連で本発明を説明したが、以上の説明から当業者には自明である通り、多くのさらなる修正、変更が可能である。例えば、本発明の圧

縮ドメインに基づく処理法を説明するために用いたブロック図は、本発明のある特定の機能および関係の動作を示すものである。これらの機能性ブロックの境界は、説明の都合上、任意に定めたものである。特定の機能および関係が適当に設けられている限り、別の境界が設定され得る。

【0091】さらに、本発明のアルゴリズムを説明するために用いた疑似コードは、構文または特定のプログラム言語を表わしてはならず、むしろ当業者が回路を組立て、または必要な処理を行うソフトウェアを作成するのに要する機能情報を提供するものである。ブロック図に示した各機能は、例えば、ソフトウェア命令によるか、あるいはデジタル信号処理回路、アプリケーション・スペシフィック集積回路（ASIC）またはこれらの組合せのような機能的に同等の回路によってインプリメントされる。以上、ここに説明した本発明では、上述のような変更、修正、適用等はすべて特許請求の範囲の精神および範疇に含まれるものとする。

【0092】

【発明の効果】完全な伸張および圧縮の工程を経由すること無しに、画像の圧縮ドメイン表現を直接操作して、選択された空間ドメイン操作を行うことができる。また、画像操作を行うのに必要な圧縮データ操作を大幅に単純化することができ、さらに品質の低下を伴わずに圧縮ドメイン内で画像データを操作することができる。

【図面の簡単な説明】

【図1】JPEG圧縮および伸張処理を示すブロック図である。

【図2】JPEGデータセットについての空間ドメインに基づく画像処理を示すブロック図である。

【図3】本発明によるJPEGデータセットについての圧縮ドメイン画像処理を示すブロック図である。

【図4】本発明による対角線フリップ（ F_d ）およびカラムフリップ（ F_y ）を用いて90°回転させたブロックを示す図である。

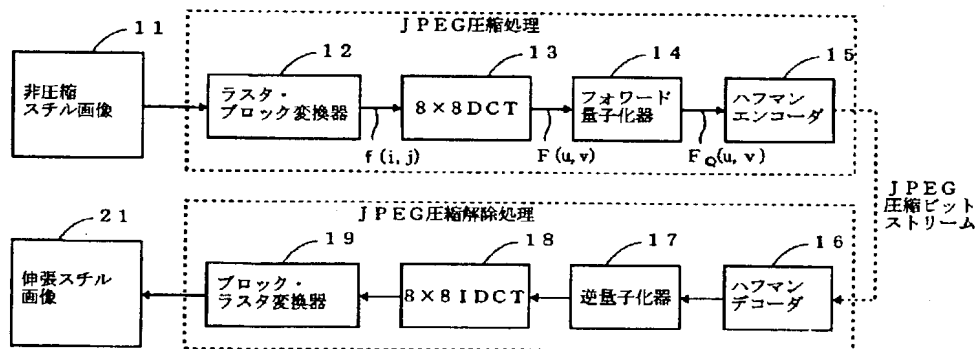
【図5】本発明との関連で用い得るデジタルスチル画像カメラ（DSC）のブロック図である。

【図6】本発明によりデジタル画像を捕捉および目視するのに用い、さらにそのような画像の処理に用いる各種構成要素間の相関関係を示すブロック図である。

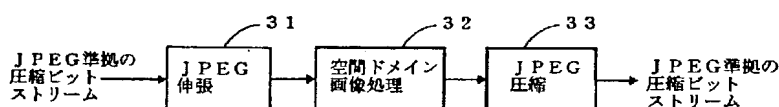
【符号の説明】

- 41 エントロピーデコーディングブロック
- 42 DCTドメイン画像処理
- 43 エントロピーコーディング
- 44 DCT係数の8×8ブロック
- 45 ジグザグ解除の手続き
- 46 DCT係数のブロックの量子化解除
- 47 ブロック指向処理
- 48 ブロック量子化
- 49 ジグザグ配置

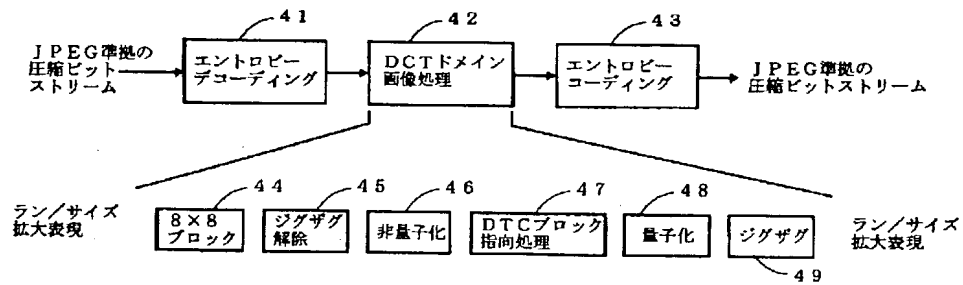
【図1】



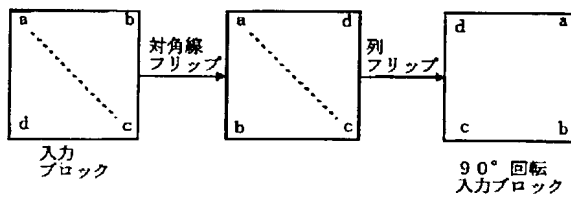
【図2】



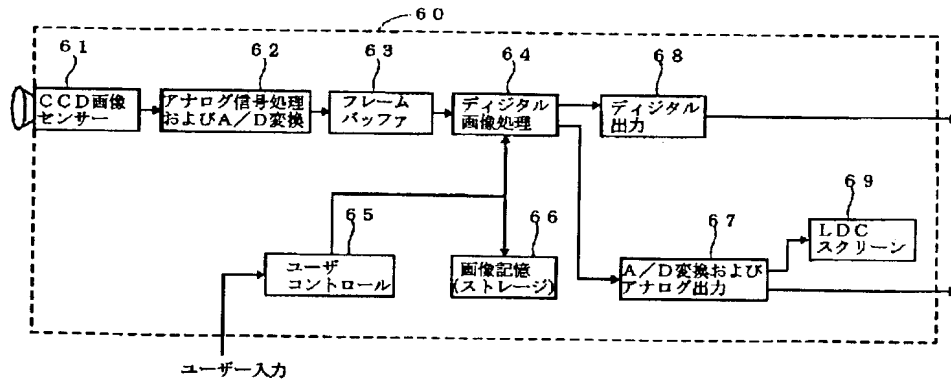
【図3】



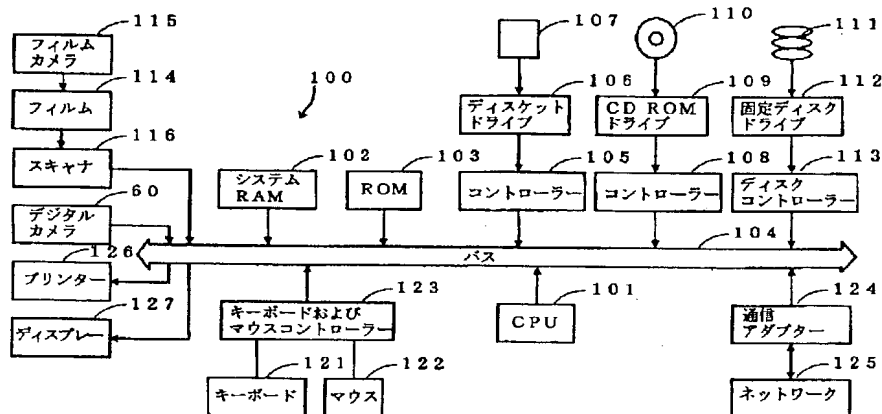
【図4】



【図5】



【図6】



フロントページの続き

(72)発明者 ヴァスデブ バスカラン
アメリカ合衆国カリフォルニア州94043
マウンテンビュー ノースウィスマンロー
ド100